

Article

Lifelong Learning from Sustainable Education: An Analysis with Eye Tracking and Data Mining Techniques

María Consuelo Sáiz Manzanares ^{1,*}, Juan José Rodríguez Díez ², Raúl Marticorena Sánchez ²,
María José Zaparain Yáñez ³ and Rebeca Cerezo Menéndez ⁴

¹ Departamento de Ciencias de la Salud, Facultad de Ciencias de la Salud, Universidad de Burgos, Research Group DATAHES, Pº Comendadores s/n, 09001 Burgos, Spain

² Departamento de Ingeniería Informática, Escuela Politécnica Superior, Universidad de Burgos, Research Group ADMIRABLE, Escuela Politécnica Superior, Avd. de Cantabria s/n, 09006 Burgos, Spain; jjrodriguez@ubu.es (J.J.R.D.); rmartico@ubu.es (R.M.S.)

³ Departamento de Historia, Geografía y Comunicación, Facultad de Humanidades y Comunicación, Universidad de Burgos, Research Group PART, Pº Comendadores s/n, 09001 Burgos, Spain; mjzaparain@ubu.es

⁴ Departamento de Psicología, Facultad de Psicología, Universidad de Oviedo, Research Group ADIR, Plaza de Feijoo, 33003 Oviedo, Asturias, Spain; cerezorebeca@uniovi.es

* Correspondence: mcsmanzanares@ubu.es; Tel.: +34-673192734

Received: 15 January 2020; Accepted: 3 March 2020; Published: 5 March 2020



Abstract: The use of learning environments that apply Advanced Learning Technologies (ALTs) and Self-Regulated Learning (SRL) is increasingly frequent. In this study, eye-tracking technology was used to analyze scan-path differences in a History of Art learning task. The study involved 36 participants (students versus university teachers with and without previous knowledge). The scan-paths were registered during the viewing of video based on SRL. Subsequently, the participants were asked to solve a crossword puzzle, and relevant vs. non-relevant Areas of Interest (AOI) were defined. Conventional statistical techniques (ANCOVA) and data mining techniques (string-edit methods and k-means clustering) were applied. The former only detected differences for the crossword puzzle. However, the latter, with the Uniform Distance model, detected the participants with the most effective scan-path. The use of this technique successfully predicted 64.9% of the variance in learning results. The contribution of this study is to analyze the teaching–learning process with resources that allow a personalized response to each learner, understanding education as a right throughout life from a sustainable perspective.

Keywords: advanced learning technologies; lifelong learning; sustainability education; eye tracking; data mining techniques

1. Introduction

1.1. Eye Tracking Metrics and Their Implications for the Analysis of Information Processing During Task Resolution

Eye-tracking technology is used as a support tool for studying human behavior in different knowledge fields (learning, marketing studies, neurological studies of various pathologies, etc.). This technological resource is used for the analysis of attention levels and relates them to the cognitive processes that a learner may employ in the course of task resolution [1]. Different types of metrics are applied by the aforesaid methodology that can be classified as either static or dynamic metrics. The first are indicators that are calculated by means of frequency analysis or measures on the data that

are recorded (fixation duration, fixation frequency, fixation times, number of fixations, gaze rate, pupil dilation, and saccades, among others). The final objective is to predict effective versus non-effective behaviors [2]. The second type of metric is obtained by recording the fixation time of the gaze within an area defined by Cartesian coordinates. That area is, in turn, divided into relevant, non-relevant, and partially relevant areas [3]. One example of this type of measurement is the scan-path metric that describes both the spatial and the temporal sequence of the fixations that the participant has completed during the completion of a task. Those indicators are thought to represent evidence that the learner has performed task-resolution processes [4].

Nevertheless, most studies that have applied the above-mentioned methodology have centered on the analysis of static gaze-fixation points and to a lesser extent on the study of ocular movement paths [5]. One reason may be that the only automatic analysis provided by eye-tracking technology is the visualization of the path that an individual follows during task completion. Nevertheless, no comparisons of the path followed by different participants or by one and the same participant at different times are completed. That information can only be obtained by processing the chains with various transformation and interpretation algorithms, such as the string-edit method applied to scan-path data. The string-edit method is based on the Levenshtein distance [6]. The objective of this technique is to differentiate, in this case, between two scan-path chains. To do so, the method is used to analyze the minimum number of operations for the transformation of one chain into another. It has been used to detect types of eye-movement errors by a participant during the resolution of a task [3]. Even though it is the most widely used method for the comparison of scan-path chains, the string-edit method has a weakness because it takes no account of the duration of the characters. Some authors [7] have therefore advised the use of algorithms that complete global alignments of two sequences. However, their use is only appropriate when the areas of interest (AOIs) have a spatial approximation. The technique was used to conduct cluster analyses [8] where the authors applied the nearest neighbor (*knn*) Machine Learning technique to determine the nearest neighbors to each cluster. In particular, the comparison of the exploratory eye-path movements of different participants during properly defined visual tasks predict repetitive and global position indices with significant correlations between the average values both for position and for sequence [8]. Other authors have likewise proposed the use of the string-edit method together with visualization techniques and cluster differentiation [9]. They concluded that dynamic metric comparison methods involve overly complex computations, both for their implementation as part of the eye-tracking methodology and for users with no previous knowledge on the matter.

1.2. Areas of Application of Eye-Tracking Technology

Learning Management Systems are currently one of the most frequent areas of application for eye-tracking technology. Relevant versus non-relevant information can be obtained from those environments [10]. As mentioned in Section 1.1, different Machine Learning techniques such as clustering are used to analyze those records. The objective is the detection of effective versus non-effective learning routes. Kurzhals and Weiskopf [11] proposed cluster maps in which different color codes are applied to determine those paths. Those authors found that each learner explored the on-screen information in a different order.

Another application area for eye-tracking technology is to confirm the effectiveness of different formats for the presentation of the information. For example, the effectiveness of using Self-Regulated Learning (SRL) has been studied during task resolution. SRL is applied by using an avatar that, through orders, guides task completion as it is presented by applying multi-channel hypermedia resources. Various studies [12,13] have found that the use of SRL unifies the scanning sequences of the participants and can even homogenize the results that relate to the variable “previous knowledge” [13]. In addition, the use of SRL improves the autonomous learning of the participant. One possible explanation is that self-regulation, together with the use of multi-channel information presentation techniques, assist adaptation to the way

in which each learner actually learns [14,15]. Lastly, those resources improve the development of the cognitive processes of selection, organization, and cognitive integration of the information [15].

Moreover, eye-tracking technology has been used for the study of reading comprehension and memory-related tasks among students. Zhai et al. [16] found that SRL techniques improved reading comprehension. They likewise found differences relating to scan-paths between men and women. Men were helped more by feedback on the task-resolution procedure, while women preferred feedback on final termination of the task.

Other investigations [17] have studied comprehension in tasks that are presented with self-regulation. Their authors found differences between the scan-path chains that could be explained by the previous vs. no previous knowledge of participants. Along those lines, the investigations [18] indicated that the previous knowledge of the participants appeared to be the cause that best explained the differences found for the resolution of written tasks but not for those presented in visual formats [19]. In addition, if the information was supported by a self-regulated video, the differences noted in the scan-paths tended to diminish [20,21], which increased learning efficacy.

Another variant in the use of eye-tracking technology was its application to instructional training, based on the visualization of task-resolution routes, prior to task completion in real spaces [22]. The results indicated that previsualization of learning routes improved task resolution in real learning environments.

1.3. Eye Tracking and Metacognitive Analysis for the Resolution of Tasks that Use Hypermedia

As mentioned in Sections 1.1 and 1.2, the use of multichannel data on SRL obtained during the use of Advanced Learning Technologies (ALTs) is a field in which eye-tracking technology is currently used more than any other, especially in Intelligent Tutoring Systems, serious games, hypermedia, and immersive virtual learning environments. Those resources appear to strengthen the use of metacognitive skills and self-regulation [23]. In this environment, Taub and Azevedo [13] studied the relation between previous knowledge and the learning results of tasks that implement SRL through a MetaTutor. The focus of those authors was on the topic of the human circulatory system. They found no significant differences in fixations on the AOIs between the participants with previous knowledge versus no previous knowledge. Nevertheless, they did find significant differences in the patterns of ocular scanning. In addition, the participants with previous knowledge were able to resolve tasks that required the application of more complex cognitive and metacognitive strategies from the participant. Likewise, Dever et al. [24] found that university students had longer fixation times on information that was presented in multimedia (text, graphic diagrams) and less so for information that was presented in written texts with no multimedia structure.

Recent investigations [25,26] have also found significant differences in the behavior of ocular scanning between expert and novice participants. The expert participants assigned their attention with greater efficiency, and they learnt more easily because they applied automated supervision processes in an autonomous manner.

Along those lines, eye-tracking technology has been used for the study of SRL in tasks that form part of serious games. Those learning spaces are designed with highly structured tasks in hierarchized sequences of difficulty [27], using techniques that both record and monitor task resolution. Recent investigations [28] have indicated that the more efficient participants accessed partially relevant or irrelevant AOIs less than those participants considered less efficient. Furthermore, the most efficient participants developed fewer sequences of Partially Relevant–Relevant and Relevant–Relevant AOIs, in comparison with the least efficient participants. Those results were considered important in the design of the ALTs. Additionally, the use of those learning spaces explained 72% of the variance in the learning results and, specifically, 62% among adults where the use of puzzle game tasks proved especially effective [29].

Likewise, these data must be placed in a global world in which education must be understood throughout the life of the individual. Likewise, the ways of teaching and learning have changed and are mediated to a high degree by technological resources (virtual platforms, hypermedia tools,

intelligent tutoring systems, etc.). In this area, teachers, especially those working with adults (Higher Education, Adult Education, etc.) must learn to teach differently, and students must learn to learn in this area. All of this will result in what has been called sustainable education, which is related to the development of personalized teaching through the use of the aforementioned resources and Educational Data Mining techniques that allow the application of algorithms to find learning patterns, and based on these, offer teaching that is as close as possible to the needs of each user. This will make both personal and material resources profitable, which supports a sustainable education system in the interest of efficiency. In summary, all the above studies agreed with the need for toolbox resources with automated processing of the registers to be incorporated in eye-tracking technology. The reasons for automated processing are explained by knowledge of the effective patterns for the resolution of different tasks and the application of prediction techniques, which permit the early detection of at-risk students and the adaptation of the learning space, in accordance with their needs [30–32]. Accordingly, technological developments that have taken place in the 21st century have also affected higher education institutions. “These changes have created an obligation to respond to the effects and needs of globalization. In response to these developments, the emphasis on lifelong learning and individuals who “learn how to learn” in higher education institutions has become more and more important in order to educate individuals to fulfill the needs of the century” [33] (p. 1). The application of resources such as eye tracking also facilitates the use of data mining techniques. These offer different forms to predict the students at risk or to classify different needs of students [34], all of which allow the teacher to make a more precise and adjusted method of teaching and therefore more sustainable.

In view of the earlier investigations, the following research questions (RQs) are proposed:

- RQ1.** Will the resolution of serious-game-based tasks depend on the characteristics of the participants (students versus teachers, with previous knowledge versus no previous knowledge, and gender)?
- RQ2.** Will the scan-path chains differ depending on the characteristics of the participants (students versus teachers, with previous knowledge versus no previous knowledge, and gender)?
- RQ3.** Will access to relevant vs. non-relevant areas of information depend on the characteristics of the participants (students versus teachers, with previous knowledge versus no previous knowledge, and gender)?
- RQ4.** Will the groupings found with the clustering technique coincide with the groups of participants (type of students versus teachers)?
- RQ5.** Will the use of string-edit methods that apply clustering analysis increase the prediction of access to relevant versus non-relevant information?

2. Materials and Methods

2.1. Participants

The sample group numbered 36 participants: 14 students and 22 university teachers (Group 1: 5 University of Experience students (>54 years old following non-regulated courses at university); Group 2: 22 university teachers; Group 3: 9 graduate and Master’s degree students at university). In Table 1, the distribution of the sample can be consulted in relation to the independent variables, gender and age (see Table 1).

Table 1. Description of the sample with regard to the variables age and gender.

Participant Type	N		Gender				
			Men		Woman		
	<i>n</i>	<i>M_{age}</i>	<i>SD_{age}</i>	<i>n</i>	<i>M_{age}</i>	<i>SD_{age}</i>	
University of experience students (Group 1)	5	4	65.25	2.87	1	68.00	-
University teachers (Group 2)	22	14	45.38	9.97	8	49	9.69
Graduate and master’s students (Group 3)	9	4	29.75	6.61	5	24.20	3.49

Note: *M_{age}* = Mean age, *SD_{age}* = Standard Deviation age.

2.2. Instruments

1. Eye tracking iViwe XTM, SMI Experimenter Center 3.0, and SMI BeGaze™ and a monitor with a resolution of 1680×1050 were used for the task resolution exercise. This tool registers ocular movements, their coordinates, and pupillary diameters of each eye. In this study, 60 Hz were applied, scan-path metrics and dynamic scan-path metrics were used, and AOI statistics were determined. The registers were analyzed with the SPSS v. 24 statistical software package.
2. A questionnaire was prepared ad hoc on the characteristics of each participant (age, gender, level of studies, present employment situation, and level of previous knowledge).
3. A crossword was prepared ad hoc on knowledge of the spoken information in the video relating to medieval monasteries.
4. A video was prepared ad hoc on the monasteries and the layout of the cloisters based on serious-games methodology with presentation of a set of 4 slides in which the information was available in summary form and supported by images, with a voice-over on the SRL-based content.

2.3. Procedure

Authorization from the University of Burgos Bioethics Committee was obtained prior to the start of the investigation. Convenience sampling was applied to select the sample. No economic compensation was offered to any of the participants. They were likewise informed of the objectives of the investigation, and in all cases their informed consent was given in writing. The investigation took place in a laboratory of the Higher Polytechnic School of Burgos University. The laboratory had two adjoining rooms. In Room 1, the participants were interviewed to obtain their personal information (age, gender, type of studies, etc. and degree of previous knowledge on the theme of the test, which was on History of Art, and in particular the architectural typology of the Medieval monastery). Subsequently, each participant entered Room 2, one by one, and the calibration test was prepared. The participant was seated, depending on the height of each person, at a distance of between 60 to 70 cm from the table. The calibration test was then performed (see Figure 1); to do so, deviation standards of 0.1–0.9 were applied to both eyes with a percentage adjustment of between 86.5–100%. Three individuals were omitted from the study as in their case they did not comply with the calibration standards. Subsequently, the test was applied that consisted of watching the video on the features of a Medieval monastery that lasted for 1:45 s. The video was designed by a specialist professor in Art History, and the voice-over by a specialist teacher in SRL. Having finished viewing the video, each participant was given a crossword puzzle with five questions on the concepts that had been covered in the video (see Figure 2). The crossword was prepared with the Crossword Puzzle application. The evaluation sessions were always directed by the same people: an expert psychologist in SRL and a computer engineer, both with expertise in the operation of eye tracking.

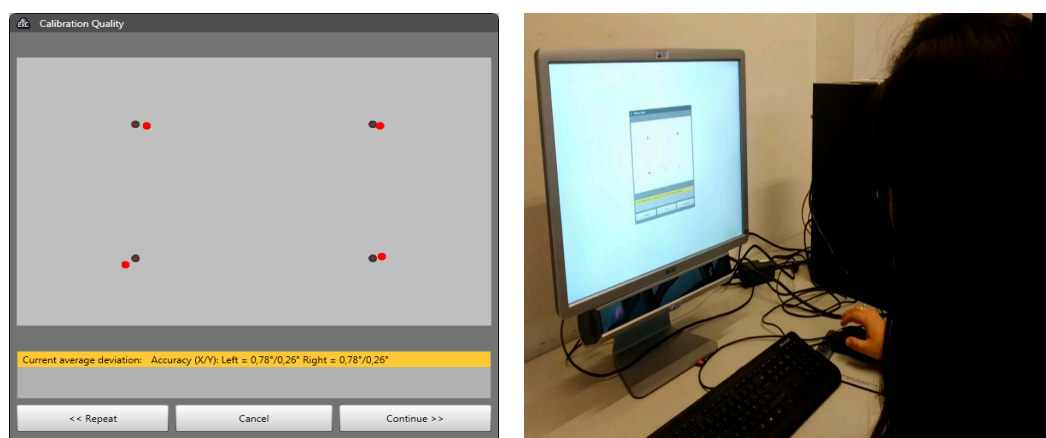


Figure 1. Calibration process with eye-tracking technology before completion of the test.

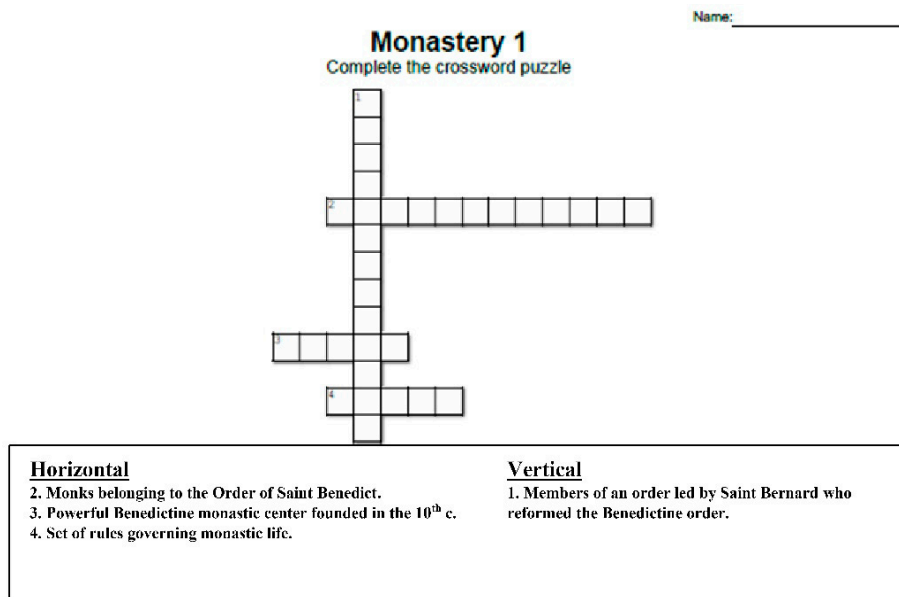


Figure 2. Crossword for completion after the test.

Regarding data preparation for the scan-path analysis, the fixations of the ocular scanning were first of all extracted. The fixation time was set at an interval of 83.1–1033.2 ms and a maximum radius of 100 pixels. The data were then transformed into data chains, and each page of the video presentation was divided up into small areas. In this case, fourteen horizontal and vertical partitions were prepared with a screen resolution of 1680 × 1050. The eye-tracking software facilitated the assignation of an alpha-numeric code in each area. An example of the spatial division of the task in the video can be examined in Figure 3. The AOIs can also be observed (Relevant area in red; Non-relevant area in green; Barely relevant area in orange). The Relevant information referred to text and image data that had been presented in the video and mentioned in the voice-over; the Non-relevant areas were spaces in white; and the Barely relevant areas referred to elements such as the avatar image or the title. A total of 36 scan-paths were analyzed by applying different string-edit methods (Uniform Distance model, City Block Distance model, and Euclidian Distance model).

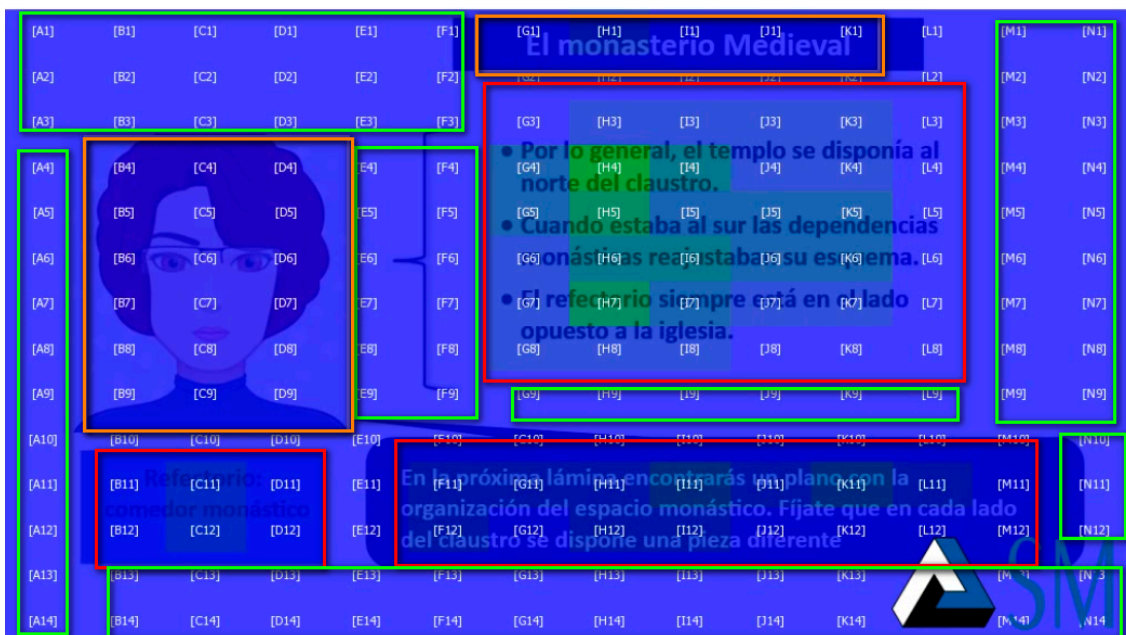


Figure 3. Geographic division of the task into quadrants and definition of the areas of interest (AOIs).

2.4. Statistical Analysis

Before commencing the study, the sample was tested to confirm that its distribution was within the parameters of normality in the variable previous knowledge. To do so, both the skewness and the kurtosis values of the selected indicators were established. The highest skewness values, |2.00|, indicate extreme skewness and the lowest values indicate that the sample follows a normal distribution [35]. Kurtosis values between |8| and |20| suggest extreme kurtosis [36]. The skewness and the kurtosis values of the other variables are also shown in Table 2 (Crossword solving, Chain length, Relevant area, Non-relevant area, and Barely relevant area). The skewness indicators were situated in an interval of |0.12| to |1.20| and those of kurtosis between |0.37| to |2.18|, which suggests no extreme deviation from normality in the sample, so parametric statistics were employed for the confirmation of the proposed hypothesis. In particular, two-factor fixed-effects ANCOVA, cluster analysis, and Cross-tables were all applied, as well as machine learning, discriminant analysis, and the string-edit method (Uniform Distance model, City Block Distance model and Euclidian Distance model).

Table 2. Skewness and kurtosis analysis of the independent variables.

	N	Minimum	Maximum	M	SD	Skewness		Kurtosis	
						S	S Error	S	S Error
Prior knowledge	36	1.00	2.00	1.53	0.51	−0.12	0.39	−2.18	0.77
Crossword solving	36	0.00	5.00	3.69	1.58	−0.88	0.39	−0.37	0.77
Chain length	36	111.00	402.00	295.89	66.16	−1.20	0.39	1.670	0.77
Relevant area	36	49	283	196.97	52.12	−1.17	0.39	1.45	0.77
No relevant area	36	25	157	79.92	27.47	0.48	0.39	0.53	0.77
Barely relevant area	36	5	44	19.00	8.59	1.03	0.39	1.89	0.77

Note. N = number of participants, M = Mean, SD = Standard Deviation, S = Statistical, S Error = Standard error.

3. Results

A two-factor fixed-effects ANCOVA (type of participant and prior knowledge) was applied to test RQ1, RQ2, and RQ3 with a covariable (gender). As can be seen in Table 3, significant differences were only found in relation to the variable “type of participant in Crossword solving” ($F_{1, 32} = 3.84, p = 0.03, 95\% \text{ CI}$). The highest averages were detected in Group 2, university teachers, both in the group with previous knowledge ($M = 5$) and in the group with no previous knowledge ($M = 3.73$). Likewise, the covariable gender only had effects for Chain length ($F_{1, 32} = 5.55, p = 0.03, 95\% \text{ CI}$). However, the value of the effect was low in both cases (20% and 16%, respectively).

An Expectation–Maximization (EM) algorithm was used to test RQ4, and the two-step cluster node (hierarchical algorithm based on BIRCH [37]) was used to determine the number of clusters with an average of 0 and a variance of 1. In all, 3 clusters were found, Cluster 1 (C1) defined as low (average between −1.0 and 0, $n = 4, 11.1\%$ of the total of participants), Cluster 2 (C2) defined as sufficient (average between 0 and 0.5, $n = 13, 36.1\%$ of the total of all participants), and Cluster 3 (C3) defined as good (average between 0.5 and 1, $n = 19, 52.8\%$ of the total of all participants) (see Table 4).

It was then studied whether the selected variables were equally sustainable in the configuration of the clusters. To do so, an ANOVA test was applied to the clusters that were found. Significant differences were found for Chain length and for both Relevant area and Non-relevant area but not for Crossword solving, or for Barely relevant area (see Table 5).

Subsequently, the Bonferroni difference of means test was applied to establish where the differences between the clusters were found (see Table 6). Significant differences were found between Cluster 1 and Cluster 2 in Chain length, the longest of which was in Cluster 2. Differences were also found between Cluster 1 and Cluster 3, the latter having the longest length. Likewise, differences were found between Cluster 2 and Cluster 3 that favored the former, and between Cluster 1 and Cluster 3 for Chain length and Relevant area in favor of Cluster 3 and in relevant and non-relevant area in favor of Cluster 1.

Table 3. Two-factor fixed effects ANCOVA (type of group, previous knowledge) and covariable gender and effect value.

	G1 N = 5		G2 N = 22		G3 N = 9		<i>F</i> _(1,32)	<i>p</i>	η^2
	1 n = 5	1 n = 7	2 n = 15	1 n = 5	2 n = 4				
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)				
Dependent variables *									
Independent variable									
Participant Type									
Crossword solving	3.60(1.51)	5(.00)	3.73(1.22)	2.60(2.30)	2.75(2.21)	3.84	0.03 *	0.20	
Chain length	300.00(19.79)	295.71(40.01)	304.80(69.59)	300.20(97.55)	252.25(95.15)	0.88	0.42	0.06	
Relevant area	69.93(9.74)	63.00(7.87)	67.24(9.57)	68.03(9.71)	60.27(13.72)	0.33	0.72	0.02	
Non-relevant area	23.78(7.11)	30.66(6.89)	26.98(9.18)	24.41(6.71)	32.79(14.78)	0.32	0.73	0.02	
Barely relevant area	6.28(3.60)	6.34(1.99)	5.78(1.52)	7.55(3.34)	6.93(2.04)	0.63	0.54	0.04	
Dependent variables *									
Independent variable							1.07	0.31	0.03
Prior knowledge									
Crossword solving							0.30	0.59	0.01
Chain length							0.19	0.66	0.01
Relevant area							0.38	0.54	0.01
Non-relevant area							0.30	0.59	0.01
Barely relevant area							1.07	0.31	0.03
Covariable sex									
Crossword solving							1.22	0.28	0.04
Chain length							5.55	0.03 *	0.16
Relevant area							0.01	0.95	0.00
Non-relevant area							0.04	0.85	0.001
Barely relevant area							1.07	0.31	0.03

* $p < 0.05$. Note: *N* = number of participants; *M* = Mean; *SD* = Standard Deviation; η^2 = eta squared (effect value); G1 = university of experience students; G2 = university teachers; G3 = graduate and master's degree students; 1 = participant with previous knowledge; 2 = participant with no previous knowledge.

Table 4. Centers of final clusters.

Variables	Clusters		
	Cluster 1 (Low)	Cluster 2 (Sufficient)	Cluster 3 (Good)
	n = 4	n = 13	n = 19
Crossword solving	3.50	3.69	3.74
Chain length	147.75	354.69	286.84
Relevant area (%)	55.69	64.47	69.45
Non-relevant area (%)	38.14	28.30	24.80
Barely relevant area (%)	6.16	7.23	5.75

Table 5. ANOVA test of final cluster membership assigned to each participant.

	Cluster		Error		<i>F</i>	<i>p</i>
	Mean Square	<i>df</i>	Mean Square	<i>df</i>		
Crossword solving	0.093	2	2.65	33	0.035	0.97
Chain length	67143.76	2	573.27	33	117.12	0.000 *
Relevant area	340.40	2	79.13	33	4.30	0.02 *
No relevant area	299.80	2	66.78	33	4.49	0.02 *
Barely relevant area	8.55	2	4.89	33	1.75	0.19

* $p < 0.05$. Note: *df* = degrees of freedom.

The next test was a Cross-table to study the relation between the groups applied to the sample from the variable "type of participant" and cluster distributions (see Table 7). In all, 80% of the participants of Group 1 were found in Cluster 3 (good), 50% in Group 2, and 44.3% in Group 3.

Table 6. Bonferroni difference of means test between the clusters for the variable type of participant.

	C1 vs. C2		C1 vs. C3		C2 vs. C3	
	SD	p	SD	p	SD	p
Crossword solving	-	-	-	-	-	-
Chain length	-206.94 *	0.000	-139.09 *	0.000	67.85 *	0.000
Relevant area	-	-	-13.75% *	0.03	-	-
Non-relevant area	-	-	13.34% *	0.02	-	-
Barely relevant area	-	-	-	-	-	-

* $p < 0.05$. Note: C1 = Cluster 1; C2 = Cluster 2; C3 = Cluster 3.

Table 7. Cross-table between cluster type and type of participant.

Type of Participant	Cluster Number of Case						Total
	1	%	2	%	3	%	
G1	0	0%	1	20%	4	80%	5
G2	2	9%	9	41%	11	50%	22
G3	2	22.3%	3	33.3%	4	44.3%	9
Total	4	11.1%	13	36.1%	19	52.7%	36

Note: G1 = university of experience Students; G2 = university teachers; G3 = graduate and master’s degree students.

A discriminant analysis of the variable “type of cluster” was then performed to see the distribution of participants assigned to each group. No significant differences were found for Wilks’ Lambda in the dependent variables, and a tendency towards the difference was only found for the variable “Crossword solving” ($p = 0.06$) (see Table 8). In Figure 4, the distribution of the participants can be seen in each of the clusters that were found.

Table 8. Wilks’ Lambda for the dependent variables.

Dependent Variables	Wilks’ Lambda	F	df ₁	df ₂	p
Crossword solving	0.84	3.10	2	33	0.06
Chain length	0.98	0.38	2	33	0.68
Relevant area	0.97	0.49	2	33	0.62
Non-relevant area	0.97	0.50	2	33	0.61
Barely relevant area	0.94	1.10	2	33	0.34

Note: *df* = degrees of freedom.

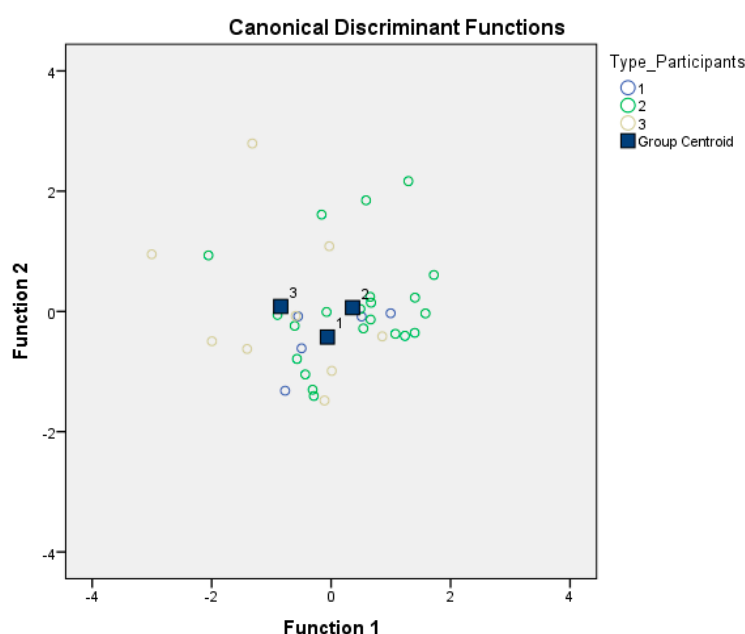


Figure 4. Canonical Discriminant Functions in relation to the variable “Cluster Type”.

Finally, an analysis of the sequences of fixations found on the scan-paths of the different participants was performed, in order to test RQ5. As already indicated in Section 2.3, a 14 × 14 grid was applied to the screen, which implies registering 196 fixation possibilities. As indicated in the section on procedure, the AOIs were classified into Relevant area (marked in red), Non-relevant area (marked in green), and Barely relevant area (marked in orange) (see Table 9).

Table 9. Definition of the Areas of Interest (AOIs).

A1	B1	C1	D1	E1	F1	G1	H1	I1	J1	K1	L1	M1	N1
A2	B2	C2	D2	E2	F2	G2	H2	I2	J2	K2	L2	M2	N2
A3	B3	C3	D3	E3	F3	G3	H3	I3	J3	K3	L3	M3	N3
A4	B4	C4	D4	E4	F4	G4	H4	I4	J4	K4	L4	M4	N4
A5	B5	C5	D5	E5	F5	G5	H5	I5	J5	K5	L5	M5	N5
A6	B6	C6	D6	E6	F6	G6	H6	I6	J6	K6	L6	M6	N6
A7	B7	C7	D7	E7	F7	G7	H7	I7	J7	K7	L7	M7	N7
A8	B8	C8	D8	E8	F8	G8	H8	I8	J8	K8	L8	M8	N8
A9	B9	C9	D9	E9	F9	G9	H9	I9	J9	K9	L9	M9	N9
A10	B10	C10	D10	E10	F10	G10	H10	I10	J10	K10	L10	M10	N10
A11	B11	C11	D11	E11	F11	G11	H11	I11	J11	K11	L11	M11	N11
A12	B12	C12	D12	E12	F12	G12	H12	I12	J12	K12	L12	M12	N12
A13	B13	C13	D13	E13	F13	G13	H13	I13	J13	K13	L13	M13	N13
A14	B14	C14	D14	E14	F14	G14	H14	I14	J14	K14	L14	M14	N14

Relevant area (marked in red), Non-relevant area (marked in green), and Barely relevant area (marked in orange).

The frequencies of the fixations on the scan-path chains completed by the different participants within the defined quadrants were found. As seen in Figure 5, the greater percentages of fixation frequencies were situated within the AOIs defined as relevant (noted in red). A clustering analysis with k-means was also performed on the fixation frequencies on the AOIs, where the points represent the clusters within the AOIs (see Figure 6).

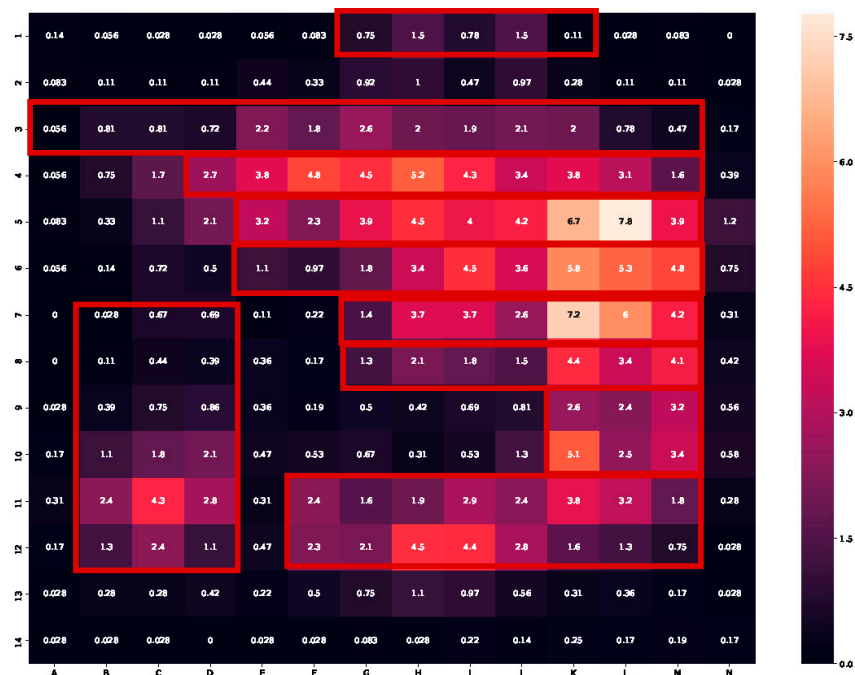


Figure 5. Frequencies within the quadrants and their relationship with the AOIs.

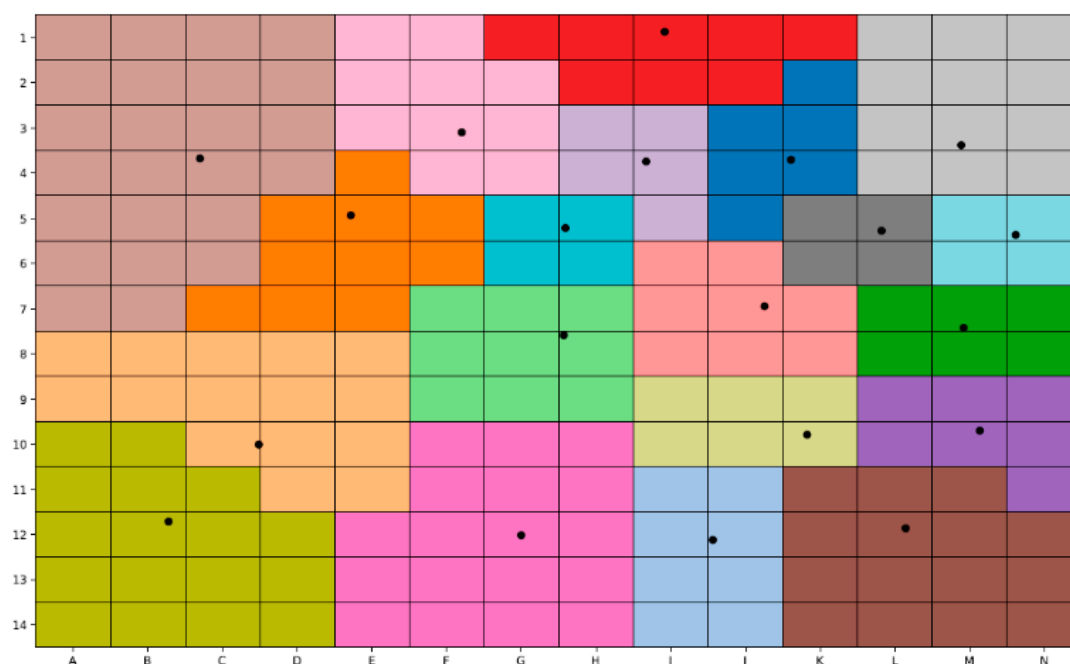


Figure 6. Analysis of clustering with k-means on the frequencies within the AOIs.

The next step was to apply three methods of analysis of the distance between the scan-path strings. In particular, three different string-edit methods were applied: (1) the Uniform Distance model, a method that assigns a cost of either 0 or 1 as a function of whether the elements of the pair are equal; (2) City Block Distance, a model that is based on the calculation of the absolute differences of the coordinates; and (3) Euclidian Distance, a model based on the Euclidean distance between coordinates. In this study, the cost of both elimination and addition was set at one, while the corresponding model was used in each case for the cost of substitution (Uniform Distance, City Block, and Euclidean Distance). The result yielded the final edition of the chains for each pair in the 36 scan-paths that were selected. The averages and the standard deviations of the distances found by each method of distance analysis and for each participant are presented below in Table 10.

Table 10. Averages and standard deviations of the distances found with each method to find the distance between each participant and the others.

Participants	Uniform Distance Model		City Block Distance Model		Euclidian Distance Model	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
AL_01	267.03	52.55	93.11	37.47	84.80	38.66
AL_02	286.53	54.17	97.49	39.79	88.11	41.05
AL_03	267.36	69.93	175.46	59.26	172.87	60.04
AL_04	279.28	73.43	203.72	64.88	200.00	65.25
AL_05	301.58	54.63	107.12	47.84	98.39	49.76
AL_06	309.78	55.52	110.06	44.23	99.75	46.21
AL_07	330.72	58.06	134.06	52.84	125.64	54.81
AL_08	265.42	52.87	88.62	38.98	80.97	39.85
AL_09	286.72	53.19	95.40	39.25	85.65	40.79
E_01	274.44	53.07	89.97	39.38	81.59	40.57
E_02	271.14	54.36	98.19	38.49	88.62	39.53
E_03	268.39	52.16	86.99	39.95	79.29	40.99
E_04	304.42	54.74	105.42	42.92	95.13	44.67
E_05	269.89	53.75	91.92	37.75	83.56	38.98
P0_01	278.47	56.42	95.26	36.83	87.39	37.93
P0_02	278.75	53.14	93.12	41.13	84.53	42.57
P0_03	296.58	53.99	100.82	47.24	92.46	48.71

Table 10. Cont.

Participants	Uniform Distance Model		City Block Distance Model		Euclidian Distance Model	
	M	SD	M	SD	M	SD
P0_06	261.47	52.89	89.85	38.40	82.43	39.41
P0_07	335.08	59.77	142.74	57.62	135.93	59.03
P0_08	262.75	61.59	129.97	43.22	125.39	43.32
P0_09	259.64	54.98	98.92	37.47	92.94	38.10
P0_10	273.06	71.11	180.09	60.27	175.67	60.50
P0_11	262.94	52.93	89.43	39.22	82.22	40.01
P0_12	324.22	57.20	126.62	51.02	117.49	53.26
P0_13	293.39	53.96	100.24	48.17	92.30	49.67
P0_14	267.17	54.14	93.14	38.87	85.58	39.63
P0_15	260.44	56.57	109.96	37.87	103.47	38.33
P0_16	305.75	54.51	106.61	49.16	98.03	50.79
P0_17	280.08	52.96	94.73	38.21	85.31	39.85
P0_18	321.61	56.35	122.78	49.62	113.45	51.92
P0_19	279.75	53.78	91.31	41.23	82.80	42.36
P0_20	275.72	54.05	95.08	38.63	85.44	39.91
P0_22	290.22	53.32	99.39	44.40	90.67	46.14
P0_23	324.78	56.95	121.18	49.58	111.66	51.61
P0_24	288.08	53.49	98.10	42.12	88.87	43.59
P0_25	257.83	55.41	100.06	38.83	93.80	39.17

Note: M = Mean; SD = Standard Deviation, AL = graduate and master’s degree students, E = university of experience students, and P0 = University professors.

The following test was to establish whether there were significant differences between the distances found with the three methods by applying a single-factor fixed-effects ANOVA (type of method). Significant differences were found both for the Uniform Distance model ($F_{1, 34} = 13006.22, p = 0.007$) and for the City Block Distance model ($F_{1, 34} = 771.23, p = 0.043$), the former model showing higher significance for discrimination of distances. In Figures 7–9, the distances between each pair found with each model are shown. The most clearly defined trajectory is the one found with the Uniform Distance model (a darker red color implies that the distance was greater between the participants (with a confidence interval of 95%).

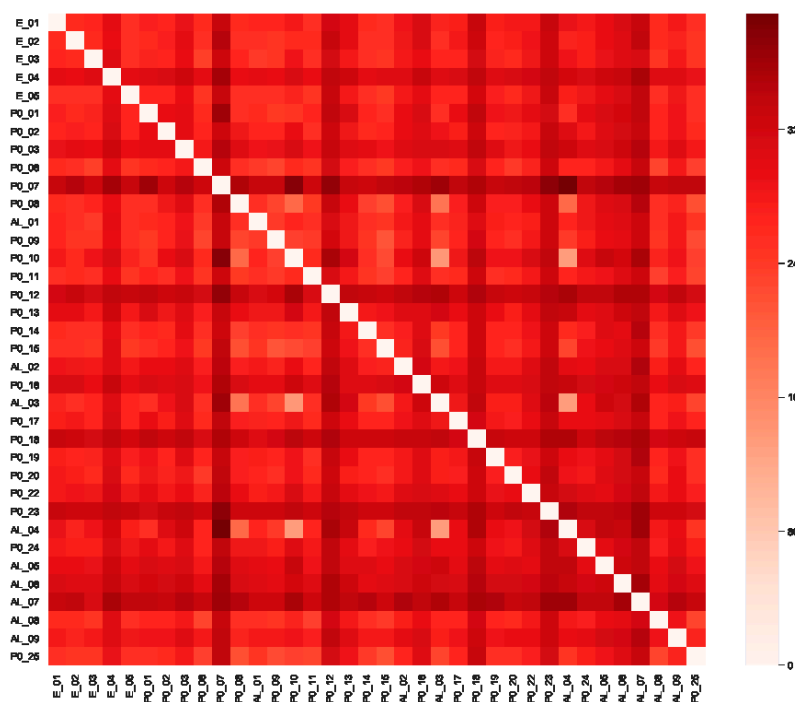


Figure 7. Heatmaps of the scan-path chains: Uniform Distance model.

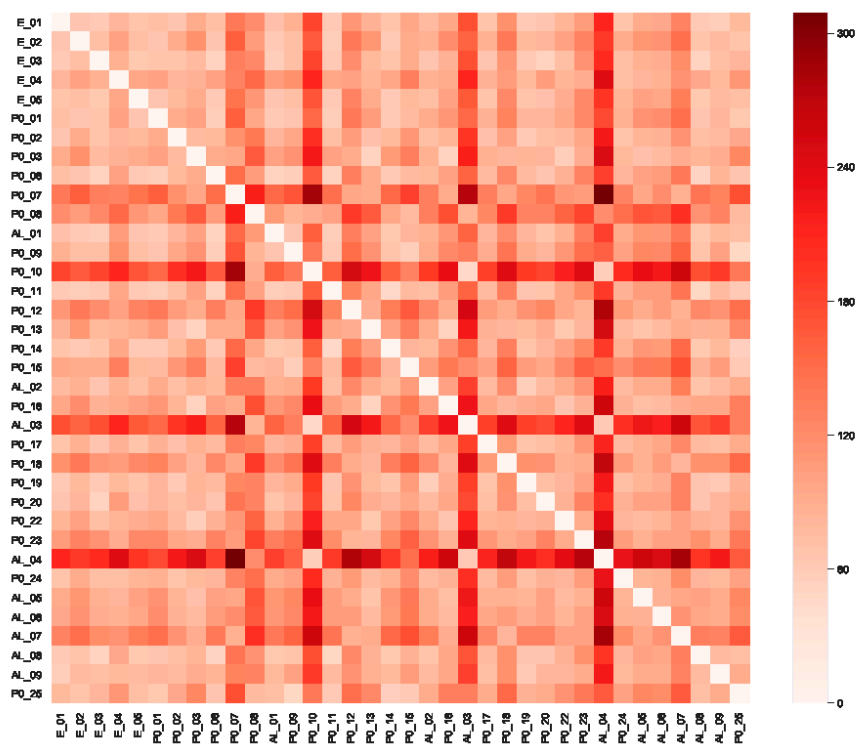


Figure 8. Heatmaps of the scan-path chains: City Block Distance model.

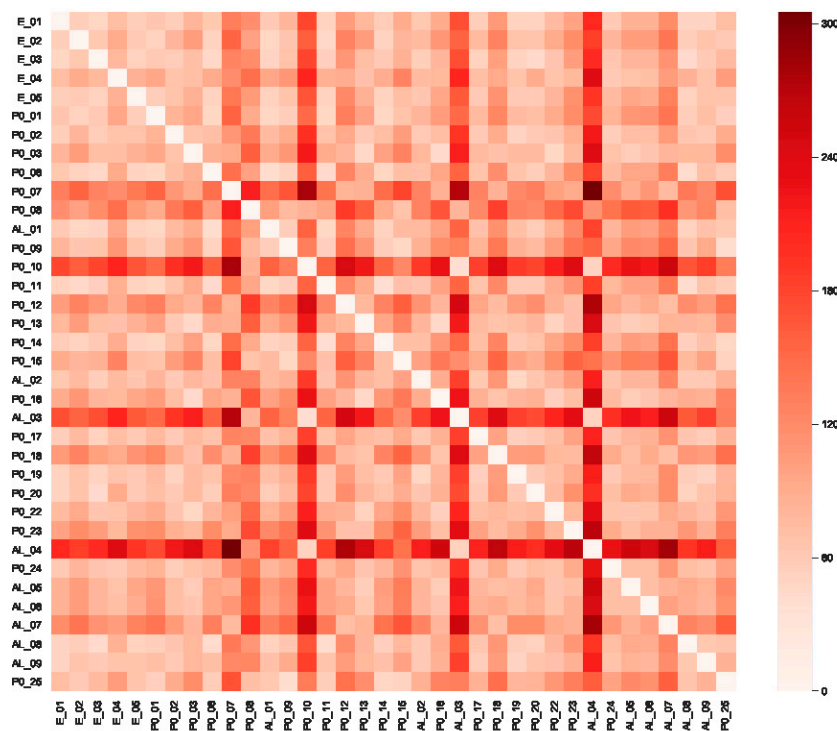


Figure 9. Heatmaps of the scan-path chains: Euclidian Distance model.

A linear regression analysis was also performed, in order to establish which of the distance analysis methods better predicted the access of the participants to relevant areas. A $R^2 = 0.691$ was found, which implies that the use of the distance analysis technique predicted 69.1% of the discrimination of the fixation of the participants within the Relevant area. In particular, the Uniform Distance model

was the only model for which significant differences were found $t = 4.82, p = 0.00$ along with a partial significant correlation ($r = 0.649, p < 0.05$).

Subsequently, multi-dimensional scaling [38] was performed with the distance models that were used to test the distribution of the distances that were obtained. As may be seen, the uniform Distance model is where a sharper grouping of the distances between the chains was observed between the chains developed by the participants where the extreme cases were the most clearly discriminated (see Figures 10–12).



Figure 10. Multi-dimensional scaling based on the distances obtained with the Uniform Distance model.



Figure 11. Multi-dimensional scaling based on the distances obtained with the City Block Distance model.



Figure 12. Multi-dimensional scaling based on the distances obtained with the Euclidian Distance model.

4. Discussion

Firstly, it must be highlighted that the use of eye-tracking technology can register data on the ocular movements of the participants during task resolution that would pass by unnoticed without this technology in an observational record [2]. In addition, dynamic scan-path registers provide data on the interaction in different positions of a sequence of action. Likewise, the application of string-edit methods facilitates discrimination and prediction of effective versus non-effective learning patterns [5]. In this area, not all the string-edit methods that apply clustering techniques appear to show greater discrimination between the distances [8,10,11]. Additionally, if this method is accompanied by visualization techniques, it will facilitate user interpretation of the results [9]. In particular, the string-edit method that was most effective in this study was the Uniform Distance model with an explained variance of 64.9%. It was also the method that best facilitated the discrimination of the scan-path followed by the at-risk learners.

Secondly, the results of this study support the findings of other investigations [12,20,26] on the way that learning tasks that apply SRL techniques, multichannel resources, and serious games produce behavioral learning patterns that are quite similar both among participants with and without previous knowledge. This form of presenting the task appears to homogenize the chain length and the results for Crossword solving [13]. Even so, the use of only statistical techniques will not lead to a precise discrimination of the differences. The application of clustering techniques has therefore highlighted that those groupings will not always coincide with the collective that is assigned a priori. In other words, from among the groups of teachers and students, there are groupings that at the same time encompass both the former and the latter, with regard to the relevant versus non-relevant AOIs and the type of chain that they use in the completion of the task. The cluster defined as the most effective follows an average not a high Chain length, and the participants have more access to relevant areas and less to irrelevant areas. Only 50% of the teachers and 44.3% of the students were found in this cluster. This fact supports the need to use string-edit methods in order to determine in a more precise manner the most effective versus non-effective path. The use of the methodology of evaluation of the learning process of each student facilitates on the one hand the recording of information that mere observation of the process does not allow. Likewise, the use of data processing techniques such as string-edit methods and cluster analysis facilitates development of the knowledge of the learning patterns of each

student and allows the grouping of students with similar patterns. These achievements are important for the design of personalized education proposals. Likewise, personalization goes hand in hand with profitability and sustainability of resource development and use [33,34].

5. Conclusions

The findings of this study open an avenue towards new investigative questions because there are other hidden variables that might be influencing the results, even though the use of ALT, SRL, and serious games appear to neutralize, in part, the influence of previous knowledge, as shown by the detection of the three clusters.

Nevertheless, the data from this study must be treated with caution, due to the characteristics of the sample, size, and sampling choice. However, it must also be taken into account that work with this methodology is laborious and implies a structure of microanalysis, which complicates the use of large sampling sizes.

In summary, further studies are needed to analyze these questions in different learning environments. In addition, fine discrimination techniques are needed that will lead to greater precision in the studies of scan-path behavior chains. Techniques that will test the findings with other characteristics of the participants may yield information on the differences and the similarities of behavioral patterns for task resolution and on the prediction of the most and the least effective patterns. These data are essential for the design of both the tasks and the ALTs, and our investigative work will continue along those same lines.

Finally, it is important to note that in the technological society, knowledge is continuously evolving. Therefore, education must be considered as a necessity throughout life. The ways of teaching and learning have also changed and are increasingly supported by technological resources. The use of techniques that allow the personalization of evaluation such as eye tracking facilitates the study of data related to the processing of information on each subject and on the other hand facilitates the use of Educational Data Mining techniques such as supervised (prediction) and unsupervised learning (grouping). In turn, the use of EDM facilitates the detection of individual and group learning patterns. Finally, these studies facilitate the detection of students at risk and the learning needs of each one of them. This detection is essential in areas of sustainable education, since the adjustment and precision of educational resources leads to a better distribution of them and the achievement of effective learning that results from increased from increased motivation and autonomy learners, all of which leads to more personalized learning in a continuous and sustainable manner.

Author Contributions: Conceptualization, M.C.S.M. and M.J.Z.Y.; methodology, M.C.S.M. and J.J.R.D.; software, J.J.R.D.; R.M.S.; R.C.M.; validation, M.C.S.M. and J.J.R.D.; formal analysis, M.C.S.M. and J.J.R.D.; investigation, M.C.S.M. and R.M.S.; resources, M.J.Z.Y. and R.M.S.; data curation, M.C.S.M.; writing—original draft preparation, M.C.S.M.; writing—review and editing, J.J.R.D., M.J.Z.Y.; R.M.S. and R.C.M.; visualization, J.J.R.D.; supervision, J.J.R.D.; project administration, M.C.S.M.; funding acquisition, M.C.S.M., M.J.Z.Y., R.C.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded through the European Project “Self-Regulated Learning in SmartArt” 2019-1-ES01-KA204-065615 and the Research Funding Program (Funding of dissemination of research results, 2020) of the Vice-Rectorate for Research and Knowledge Transfer of the University of Burgos to the Recognized Investigation Group DATAHES.

Acknowledgments: To all the people and entities that have collaborated in this study, especially to the Research Group ADIR of Universidad de Oviedo for loaning the equipment Eye tracking iViwe XTM, SMI Experiment Center 3.0 and SMI BeGazeTM and the Director of the Experience University of the University of Burgos.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Ethics Statement: The Ethics Committee of the University of Burgos approved this study N° IR20/2018. Written informed consent was in each case requested from the parents and, where applicable, the legal guardians of the participating students. They all gave their written informed consent in accordance with the Declaration of Helsinki.

Data Availability Statement: The datasets generated for this study are available on request to the corresponding author.

References

1. Joe Louis Paul, I.; Sasirekha, S.; Uma Maheswari, S.; Ajith, K.A.M.; Arjun, S.M.; Athesh Kumar, S. Eye gaze tracking-based adaptive e-learning for enhancing teaching and learning in virtual classrooms. In *Information and Communication Technology for Competitive Strategies, Proceedings of the Information and Communication Technology for Competitive Strategies, Singapore, 16–17 December 2017*; Fong, S., Akashe, S., Mahalle, P.N., Eds.; Springer: Singapore, 2019; pp. 165–176. [[CrossRef](#)]
2. Steichen, B.; Carenini, G.; Conati, C. User-adaptive information visualization—Using eye gaze data to infer visualization tasks and user cognitive abilities. In *Proceedings of the International Conference on Intelligent User Interfaces, Santa Monica, CA, USA, 19–22 March 2013*; pp. 317–328. [[CrossRef](#)]
3. Takeuchi, H.; Matsuda, N. Scan-path analysis by the string-edit method considering fixation duration. In *Proceedings of the 6th International Conference on Soft Computing and Intelligent Systems, and the 13th International Symposium on Advanced Intelligence Systems, Kobe, Japan, 20–24 November 2012*; pp. 1724–1748. [[CrossRef](#)]
4. Drusch, G.; Bastien, J.M.C. Analyzing Web pages visual scanpaths: Between and within tasks variability. *Work* **2012**, *41* (Suppl. 1), 1559–1566. [[CrossRef](#)] [[PubMed](#)]
5. Xia, C.; Han, J.; Qi, F.; Shi, G. Predicting human saccadic scanpaths based on iterative representation learning. *IEEE Trans. Image Process.* **2019**, *28*, 3502–3515. [[CrossRef](#)] [[PubMed](#)]
6. Levenshtein, V.I. Binary codes capable of correcting deletions, insertions and reversals. *Sov. Phys. Dokl.* **1966**, *10*, 707–710.
7. West, J.M.; Haake, A.R.; Rozanski, E.P.; Karn, K.S. EyePatterns: Software for identifying patterns and similarities across fixation sequences. In *Proceedings of the Conference of Eye Tracking Research & Application Symposium, San Diego, CA, USA, 27–29 March 2006*; pp. 149–154. [[CrossRef](#)]
8. Duchowski, A.T.; Driver, J.; Jolaoso, S.; Tan, W.; Ramey, B.N.; Robbins, A. Scanpath comparison revisited. In *Proceedings of the Conference of the Symposium on Eye-Tracking Research & Applications, Austin, TX, USA, 22–24 March 2010*; pp. 219–226. [[CrossRef](#)]
9. Raschke, M.; Herr, D.; Blascheck, T.; Ertl, T.; Burch, M.; Willmann, S.; Schrauf, M. A visual approach for scan path comparison. In *Proceedings of the Eye Tracking Research and Applications Symposium (ETRA), New York, NY, USA, 12–13 March 2014*; pp. 135–142. [[CrossRef](#)]
10. Khedher, A.B.; Jraidi, I.; Frasson, C. Tracking students' analytical reasoning using visual scan paths. In *Proceedings of the IEEE 17th International Conference on Advanced Learning Technologies (ICALT), Timisoara, Romania, 3–7 July 2017*; pp. 53–54. [[CrossRef](#)]
11. Kurzhals, K.; Weiskopf, D. Space-time visual analytics of eye-tracking data for dynamic stimuli. *IEEE Trans. Vis. Comput. Graphics* **2013**, *19*, 2129–2138. [[CrossRef](#)] [[PubMed](#)]
12. Sáiz-Manzanares, M.C.; Zaparaín, M.J.; Marticorena, R.; Velasco, R. *Task Analysis with Eye Tracking Technology: SRL in SmartArt*; Peralbo, M., Ed.; Servicio de Publicaciones de la Universidad de La Coruña: La Coruña, Spain, 2019; pp. 4093–4104. [[CrossRef](#)]
13. Taub, M.; Azevedo, R. How does prior knowledge influence eye fixations and sequences of cognitive and metacognitive SRL processes during learning with an intelligent tutoring system? *Int. J. Artif. Intell. Educ.* **2019**, *29*, 1–28. [[CrossRef](#)]
14. Mayer, R. *The Cambridge Handbook of Multimedia Learning*; Cambridge University Press: Cambridge, UK, 2014.
15. Mayer, R.E. Using multimedia for e-learning: Multimedia for e-learning. *J. Comput. Assist. Learn.* **2017**, *33*, 403–423. [[CrossRef](#)]
16. Zhai, X.; Fang, Q.; Dong, Y.; Wei, Z.; Yuan, J.; Cacciolatti, L.; Yang, Y. The effects of biofeedback-based stimulated recall on self-regulated online learning: A gender and cognitive taxonomy perspective. *J. Comput. Assist. Learn.* **2018**, *34*, 775–786. [[CrossRef](#)]
17. Krejtz, K.; Duchowski, A.T.; Kopacz, A. Gaze transitions when learning with multimedia. *J. Eye Mov. Res.* **2016**, *9*, 1–17. [[CrossRef](#)]
18. Yang, F.-Y.; Chang, C.-Y.; Chien, W.-R.; Chien, Y.-T.; Tseng, Y.-H. Tracking learners' visual attention during a multimedia presentation in a real classroom. *Comput. Educ.* **2013**, *62*, 208–220. [[CrossRef](#)]
19. Chen, S.-C.; Hsiao, M.-S.; She, H.-C. The effects of static versus dynamic 3D representations on 10th grade students' atomic orbital mental model construction. *Comput. Hum. Behav.* **2015**, *53*, 169–180. [[CrossRef](#)]

20. Ho, H.N.J.; Tsai, M.J.; Wang, C.Y.; Tsai, C.C. Prior knowledge and online inquiry-based science reading: Evidence from eye tracking. *Int. J. Sci. Math. Educ.* **2014**, *12*, 525–554. [[CrossRef](#)]
21. Stull, A.T.; Fiorella, L.; Mayer, R.E. An eye-tracking analysis of instructor presence in video lectures. *Comput. Hum. Behav.* **2018**, *88*, 263–272. [[CrossRef](#)]
22. Seifert, L.; Cordier, R.; Orth, D.; Courtine, Y.; Croft, J.L. Role of route previewing strategies on climbing fluency and exploratory movements. *PLoS ONE* **2017**, *12*, e0176306. [[CrossRef](#)] [[PubMed](#)]
23. Azevedo, R.; Gašević, D. Analyzing multimodal multichannel data about self-regulated learning with advanced learning technologies: Issues and challenges. *Comput. Hum. Behav.* **2019**, *96*, 207–210. [[CrossRef](#)]
24. Dever, D.A.; Wiedbusch, M.; Azevedo, R. Learners' gaze behaviors and metacognitive judgments with an agent-based multimedia environment. In *Artificial Intelligence in Education*; Isotani, S., Millán, E., Ogan, A., Hastings, P., McLaren, B., Luckin, R., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 58–61.
25. Bruder, C.; Hasse, C. Differences between experts and novices in the monitoring of automated systems. *Int. J. Ind. Ergon.* **2019**, *72*, 1–11. [[CrossRef](#)]
26. Takacs, Z.K.; Bus, A.G. How pictures in picture storybooks support young children's story comprehension: An eye-tracking experiment. *J. Exp. Child Psychol.* **2018**, *174*, 1–12. [[CrossRef](#)]
27. Taub, M.; Azevedo, R.; Mudrick, N.; Clodfelter, E.; Bouchet, F. Can scaffolds from pedagogical agents influence effective completion of sub-goals during learning with a multi-agent hypermedia-learning environment? In Proceedings of the International Conference Learn Sciences (ICLS), Boulder, CO, USA, 23–27 June 2014; pp. 1052–1056.
28. Taub, M.; Azevedo, R.; Bradbury, A.E.; Millar, G.C.; Lester, J. Using sequence mining to reveal the efficiency in scientific reasoning during STEM learning with a game-based learning environment. *Learn. Instr.* **2018**, *54*, 93–103. [[CrossRef](#)]
29. Gallego Durán, F.; Villagrà-Arnedo, C.; Satorre Cuerda, R.; Compañ Rosique, P.; Molina Carmona, R.; Llorens Largo, F. Panorámica: Serious games, gamification y mucho más. *ReVisión* **2014**, *7*, 13–23.
30. Alemdag, E.; Cagiltay, K. A systematic review of eye tracking research on multimedia learning. *Comput. Educ.* **2018**, *125*, 413–428. [[CrossRef](#)]
31. Eitel, A. How repeated studying and testing affects multimedia learning: Evidence for adaptation to task demands. *Learn. Instr.* **2016**, *41*, 70–84. [[CrossRef](#)]
32. Gutiérrez, J.; David, E.; Rai, Y.; Le Callet, P. Toolbox and dataset for the development of saliency and scanpath models for omnidirectional/360° still images. *Signal Process. Image Commun* **2018**, *69*, 35–42. [[CrossRef](#)]
33. Taşçı, G.; Titrek, O. Evaluation of Lifelong Learning Centers in Higher Education: A Sustainable Leadership Perspective. *Sustainability* **2020**, *12*, 22. [[CrossRef](#)]
34. Aljohani, N.R.; Fayoumi, A.; Hassan, S. Predicting at-risk students using clickstream data in the virtual learning environment. *Sustainability* **2019**, *11*, 7238. [[CrossRef](#)]
35. Bandalos, D.L.; Finney, S.J. Item parceling issues in structural equation modeling. In *New Developments and Techniques in Structural Equation Modeling*; Marcoulides, G.A., Schumacker, R.E., Eds.; Lawrence Erlbaum Associates Publishers: Mahwah, NJ, USA, 2001; pp. 269–296.
36. Arias, B.; Verdugo, M.Á.; Navas, P.; Gómez, L.E. Factor structure of the construct of adaptive behavior in children with and without intellectual disability. *Int. J. Clin. Health Psychol.* **2013**, *13*, 155–166. [[CrossRef](#)]
37. Zhang, T.; Ramakrishnan, R.; Livny, M. BIRCH: An efficient data clustering method for very large databases. In Proceedings of the ACM SIGMOD International Conference on Management of Data, Montreal, QC, Canada, 4–6 June 1996; pp. 103–114. [[CrossRef](#)]
38. Borg, I.; Groenen, P.J.F. *Modern Multidimensional Scaling: Theory and Applications*, 2nd ed.; Springer: New York, NY, USA, 2005.

